

빅데이터 시스템 용어정리 과제

컴퓨터공학과 20154215 구명희

1. 에지 컴퓨팅(Edge Computing)

- >> 에지 컴퓨팅은 사용자와 가까운 위치에서 서비스를 제공하는 것을 의미합니다.
- 데이터 센터에서 서비스를 제공받는게 아닌 사용자와 가까운 어딘가에서 서비스를 제공받는 것이 목표이며, 컴퓨팅 능력을 가졌고, 네트워크 라우터 등 서비스 요청이 지나가는 길이라면 에지 컴퓨팅을 제공할 수 있는 후보가 될 수 있습니다. 기지국이나 기존 네트워크 인프라에 해당하는 스위치나 라우터, 또는 공유기나 가정용 셋톱박스 등도 에지 컴퓨팅을 제공할 수 있습니다.
 - 에지 컴퓨팅 서비스 지점을 에지라고 부르며, 에지 컴퓨팅은 사용자 가까이 위치한 에지에서 컴퓨팅 능력을 통해 서비스를 제공하는 걸 목표로 합니다.
 - 매년 유망한 IT에 관련된 시장 분야를 분석해서 리포트 형식으로 제공하고 있는 가트너는 2018년 전략 기술중 하나로 에지 컴퓨팅을 선정했습니다. 몇 년 전부터 크게 화제가 되고있는 머신 러닝과 블록 체인등과 함께 선정된걸 고려하면 에지 컴퓨팅 역시 미래에 주목받는 핵심 기술이 될 듯 합니다.
 - https://m.blog.naver.com/PostView.nhn?blogId=alice_k106&logNo=221300292775&proxyReferer=https%3A%2F%2Fwww.google.com%2F

2. 사물 인터넷(IoT, Internet of Things)

- >> 각종 사물에 센서와 통신 기능을 내장하여 인터넷에 연결하는 기술입니다.
- 사물 인터넷(Internet Of Things, IOT)은 임베디드 시스템(아두이노 등) 또는 여러 하드웨어나 센서들을 인터넷과 연결해서 사용합니다. 무선 통신을 통해 각종 사물을 연결하는 기술이며, 인터넷으로 연결된 사물들이 데이터를 주고받아 스스로 분석하고 학습한 정보를 사용자에게 제공하거나 사용자가 이를 원격 조정할 수 있는 인공지능 기술입니다.
 - IOT라는 용어는 1999년, MIT의 Auto-ID center였던 Procter & Gamble의 Kevin Ashton에 의해 만들어졌습니다.
 - IOT는 유선으로 연결되거나 블루투스, WiFi, ZigBee등의 모듈을 사용해서 무선으로 연결될 수 있습니다. 또한, 최근에는 인터넷과 연결해서 제어할 수 있는 IOT 장치가 크게 성장하고 있으며, 2020년에는 약 300억개의 장치가 생산될 것으로 추정됩니다.
 - <https://ko.wikipedia.org/wiki/%EC%82%AC%EB%AC%BC%EC%9D%B8%ED%84%B0%EB%84%B7>

3. 디바이스 드라이버(Device Driver)

- >> 특정 하드웨어나 장치를 제어하기 위한 커널의 일부분으로 동작하는 프로그램입니다.
- 높은 수준의 컴퓨터 프로그램들이 컴퓨터 H/W장치와 상호작용하기 위해 만들어진 하나의 컴퓨터 프로그램입니다.
 - https://ko.wikipedia.org/wiki/%EC%9E%A5%EC%B9%98_%EB%93%9C%EB%9D%BC%EC%9D%B4%EB%B2%84

4. 밴드위드(Band Width)

- >> BandWidth, 대역폭은 신호처리와 컴퓨팅 두가지 의미를 가집니다.
- 신호처리에서 BandWidth, 대역폭은 특정한 기능을 수행할 수 있는 주파수의 범위를 말합니다.
헤르츠 단위로 측정되며, 통신 시스템의 자료 전송율이나 주파수 범위 또는 둘 다를 말합니다.
[https://ko.wikipedia.org/wiki/%EB%8C%80%EC%97%AD%ED%8F%AD_\(%EC%8B%A0%ED%98%B8_%EC%B2%98%EB%A6%AC\)](https://ko.wikipedia.org/wiki/%EB%8C%80%EC%97%AD%ED%8F%AD_(%EC%8B%A0%ED%98%B8_%EC%B2%98%EB%A6%AC))
 - 컴퓨팅에서 대역폭은 일반적으로 초당 여러비트로 표현되는 사용가능하거나 소비된 정보용적의 비트레이트입니다.
네트워크 대역폭 또는 데이터 대역폭, 디지털 대역폭이라고도 합니다.
[https://ko.wikipedia.org/wiki/%EB%8C%80%EC%97%AD%ED%8F%AD_\(%EC%BB%B4%ED%93%A8%ED%8C%85\)](https://ko.wikipedia.org/wiki/%EB%8C%80%EC%97%AD%ED%8F%AD_(%EC%BB%B4%ED%93%A8%ED%8C%85))

5. 국가 데이터센터 위치

- >> 우리나라 데이터센터는 50여개 이상의 데이터센터가 운영되고 있습니다.
- 네이버, 삼성, KT, LGU+, SK BroadBand, LGCNS, 포스코ICT, 농협, 신한, 현대, 대한항공, 한화, 교보 등
 - 데이터 센터의 대부분은 수도권에 위치하고 있으며, 이는 운영 인력과 방문 인원의 접근성 및 인프라의 유리함에 기반을 두고 있습니다.

- 데이터 센터의 위치는 [서울 17개, 성남 5개, 용인 5개, 인천 4개, 춘천 2개, 부산 2개, 기타 8개 기타지역 ; 안양, 수원, 과천, 의왕, 하남, 파주, 고양, 대구] 와 같습니다.
- <https://blog.naver.com/0330community/221503659458>

6. 스마트 그리드(Smart Grid)

- >> 전기가 흐르는 모든 것을 둑어서 효율적으로 관리하는 신개념 시스템입니다.
- 스마트 그리드의 핵심은 전력망에 정보통신기술을 합쳐서 소비자와 전력회사가 실시간으로 정보를 주고받는 것에 있습니다. 따라서 소비자는 집, 사무실, 공장, 어느 곳에서나 사용한 전기요금을 실시간으로 확인할 수 있으며, 전기요금이 비싼 낮 시간대를 피해서 전기를 사용하는 것도 가능합니다.
 - https://ko.wikipedia.org/wiki/%EC%8A%A4%EB%A7%88%ED%8A%B8_%EA%B7%B8%EB%A6%AC%EB%93%9C

7. 정형 데이터 vs 비정형 데이터

- >> 데이터의 형태에 따라 정형 데이터와 비정형 데이터로 나눌 수 있습니다.
- 형태가 있고 연산이 가능하면 그것은 정형 데이터입니다.
정형데이터의 예로는 관계형 데이터베이스와 스프레드시트, CSV등이 있으며, 데이터 수집 난이도가 낮습니다. 이는 내부 시스템인 경우가 대부분이기 때문이며, 정형 데이터는 주로 관계형 데이터베이스에 저장됩니다.
 - 반대로 비정형 데이터는 구조화 된 데이터가 아닌, 연산이 불가능한 데이터입니다.
비정형 데이터의 예로는 SNS 데이터(트위터, 페이스북 등), 영상, 이미지, 음성, 텍스트(PDF 등) 등이 있으며, 데이터 수집 난이도가 높습니다. 이는 텍스트 마이닝 혹은 파일일 경우에 파일을 데이터 형태로 파싱해야 하기 때문이며, 비정형 데이터는 주로 NoSQL에 저장됩니다.
 - <https://beji.tistory.com/entry/%EC%A0%95%ED%98%95-%EB%B0%98%EC%A0%95%ED%98%95-%EB%B9%84%EC%A0%95%ED%98%95-%EB%8D%B0%EC%9D%B4%ED%84%B0>